

Instructions for GO Analysis

GO Analysis Overview

Background

- This Panther GO website will do GO analysis for you (<https://www.pantherdb.org>)
- GO looks at the list of genes that are significantly different, and compares it to full list of genes that were measured in the experiment. MEASURED is important. It is a mistake to compare to *all* genes in the genome. Not all genes are expressed in your cell. You need to compare to just the genes actually *in* your experiment.
- I already had you make these lists! These should be in a folder on your local computer. Only you know where you put them, but the path should be in the format “<your/computer/path>/day9/results”

General Protocol

1. Go to <https://www.pantherdb.org>
2. On this page there are numbered box.
3. In box 1, there is a “Choose File” button. This is where you upload your list of significantly differentially expressed genes. Do that now.
4. In box 2, select “Homo sapiens”
5. In step 3, select “Statistical overrepresentation test”
 - a. Then choose “GO biological processes complete”
6. Now hit submit.

The screenshot shows the Panther GO analysis web interface. It is divided into three main sections, each with a numbered box highlighting a step from the protocol:

- Box 1: Upload IDs** - This section contains a text input field for "Enter IDs:", a "Choose File" button, and a "File format" dropdown. The "Choose File" button is highlighted with a purple box. Below it, there are radio buttons for "ID List" (selected), "Previously exported text search results", "Workspace list", "PANTHER Generic Mapping", and "ID's from Reference Proteome Genome". A dropdown menu for "Organism for id list" is set to "Abrus precatorius (ABRPR)".
- Box 2: Select organism** - This section contains a dropdown menu for selecting an organism. The list shows "Homo sapiens", "Mus musculus", "Rattus norvegicus", "Gallus gallus", and "Danio rerio". "Homo sapiens" is selected and highlighted with a purple box.
- Box 3: Select Analysis** - This section contains radio buttons for "Functional classification viewed in gene list" and "Functional classification viewed in graphic charts". Below these, there are radio buttons for "Statistical overrepresentation test" (selected) and "GO biological process complete". The "Statistical overrepresentation test" option is highlighted with a purple box.

At the bottom of the form, there is a "submit" button highlighted with a purple box.

You have been redirected to a new webpage. This is where you upload your “reference”, also known as your “background”, list.

7. At the bottom of the box there is a “Choose File” button. Click that and select your background list of genes.
8. You need to then hit the blue “Upload list” button. *You can ignore the box on the right for this exercise.*

SELECT REFERENCE LIST ?

For a reference list, you may upload your own list (recommended) or choose from available whole genome

Upload Reference List from flat file or Workspace

Select Organism: (Not applicable for Generic mapping file or Reference Proteome ids)

Homo sapiens
Mus musculus
Rattus norvegicus
Gallus gallus
Danio rerio

Upload list:
Please select list type...

☒ Gene, Transcript, Protein and Alternate ID
☐ PANTHER Generic Mapping File
☐ ID's from Reference Proteome Genome

Organism for id list: **Abrus precatorius (ABRPR)**

☐ VCF file
☐ Flanking region
20 Kb

Upload list:

genes_p53peak.txt
[supported IDs](#)

Upload list

If there are redundant IDs, only the first will be used in the analysis.

Please [login](#) to be able to select lists from your workspace.

It will take seconds-to-minutes to run, but then you will get results!!! They should look like the screenshot below. It worked! We got p53 back!

Congrats, now you know how to do GO analysis.

Results ?

	Reference list	genes_sig_p53peak.txt
Uniquely Mapped IDs:	1894 out of 1934	74 out of 74
Unmapped IDs:	218	9
Multiple mapping information:	50	0

Export [Table](#) [XML with user input ids](#) [JSON with user input ids](#)

Displaying only results for FDR P < 0.05, [click here to display all results](#)

	genes_p53peak.txt (REF)	genes_sig_p53peak.txt (Hierarchy NEW! ?)
GO biological process complete	#	# expected Fold Enrichment +/- raw P value FDR
intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator	9	6 .34 17.42 + 1.97E-07 8.63E-04
↳ intrinsic apoptotic signaling pathway by p53 class mediator	14	6 .54 11.20 + 6.04E-06 1.76E-02
↳ signal transduction by p53 class mediator	24	9 .92 9.80 + 8.88E-08 7.78E-04